Information Control in the Hold-up Problem*

Anh Nguyen[†] Teck Yong Tan[‡]

February 2018

Abstract

We study the use of information control to mitigate hold-up risks. Our main result identifies a separation between information that creates *ex-ante investment incentive* and information that causes *ex-post inefficiency*, which then allows ex-post inefficiency to be eliminated without compromising the ex-ante investment incentive. We characterize the properties of the optimal information structure and the investment levels and welfare achievable with information control in the presence of hold-up risks.

Keywords: Hold-up; Information control
JEL Classification: D42, D82, D83.

*We thank Navin Kartik and Ran Spiegler for helpful comments.
[†]Department of Economics, Columbia University. Email: anh.nguyen@columbia.edu.
[‡]Department of Economics, Nanyang Technological University. Email: teckyongtan@ntu.edu.sg.

1 Introduction

This paper revisits the classic hold-up problem which arises whenever economic agents have to make sunk investments that are ex-ante uncontractible. Hold-up risks are ubiquitous in many economic settings. These include relationship-specific investment in partnerships, acquisition of firm-specific skills by employees, provision of general training by firms, campaign contribution in political lobbying, and quality investment by hospitals in the healthcare market. Framing it in the context of a bilateral trade, a buyer (he) can make an investment to increase his valuation for a good sold by a monopolist seller (she); but anticipating that the seller will charge a higher price upon investment and expropriate all the gains, the buyer never invests.

Some papers (e.g. Gul, 2001; Lau, 2008) have suggested that hiding the investment from the seller could mitigate hold-up risks. The resulting asymmetric information limits the seller's ability to extract the investment gains, thus improving *ex-ante efficiency* by partially restoring the buyer's investment incentive. However, the asymmetric information also creates the possibility of trade breaking down, thus resulting in *ex-post inefficiency.* This suggests that a tradeoff exists between creating ex-ante investment incentive and minimizing ex-post inefficiency, thus the optimal information control must balance the two effects.

In this paper, we study the use of information control to mitigate hold-up risks. A signal structure, which is publicly determined before investment, generates signals about the buyer's eventual valuation of the good. Our main result illustrates that the tradeoff described above is not necessary, if there is access to a slightly richer form of information control. Intuitively, ex-ante investment incentive is created when the seller is unaware of the buyer's higher valuation at least some of the time so that he can reap some investment gain from being under-charged; this only concerns hiding information from the seller in the "investment state". On the other hand, ex-post inefficiency is eliminated by revealing the buyer's low valuation so that the seller does not set a high price which prohibits trade; this only concerns information about the "non-investment state" which does not affect the ex-ante investment incentive. In turn, this separation implies that ex-post inefficiency can be eliminated without compromising the buyer's ex-ante investment incentive.

Our goal is to formalize this intuition which can be instructive for information design to mitigate hold-up risks in various applied contexts. To achieve this, we will begin our analysis with a hold-up setting with a single and deterministic investment in Section 3. This simple setting, whereby the investment outcome is binary, allows for a direct and transparent exposition but manages to illustrate the key intuitions. We then generalize the results to allow for stochastic investment (Section 4) and multiple types of investment (Section 5).

In the single investment setting, we first show that when the seller cannot directly observe the investment, the buyer's investment decision must be randomized in equilibrium. This is because if the seller anticipates that the buyer always invests, she will charge a high price which destroys the buyer's ex-ante investment incentive. On the other hand, the seller will charge a low price if she anticipates that the buyer never invests; in turn, the buyer will want to invest for his own gains. We then characterize the set of possible investment probability and social welfare that can be sustained in equilibrium, and the signal structure that implements it. We show that every implementable investment probability can be optimally (in terms of the social welfare) implemented by the same signal structure which is unique within an appropriate class of signal structures. This optimal signal structure takes the following form: the buyer's low valuation generates the "low" signal all the time; whereas his high valuation generates both "low" and " high" signals with strictly positive probability. The simplicity of the optimal signal structure also allows it to be replicated by practical arrangements, thus also alleviating the usual concern in the information design literature about how one derives the ability to commit to a signal structure.

The optimal signal structure can be understood by viewing the signal structure as a hypothesis test for the presence of the investment. If the test is sufficiently accurate, the seller will set the high price when the test detects an investment, and she will set the low price otherwise. Therefore, when the test makes a "false positive" type I error (i.e. detect an investment when there is not), the seller will set the high price when the buyer's valuation is low, thus leading to no trade. On the other hand, when the test makes a "false negative" type II error (i.e. fail to detect the investment), the seller will set the low price when the buyer's valuation is high, thus leaving some surplus for the buyer to compensate for his investment. Therefore, from an ex-ante perspective, ex-post inefficiency due to no trade is caused by type I errors, whereas the ex-ante investment incentive for the buyer is created from the possibility of type II errors. Since the two types of errors are created separately, the welfare maximizing hypothesis test (or signal structure) eliminates the type I errors while maintaining the presence of type II errors, thus leading to the form of the optimal signal structure described above. However, the need to maintain the accuracy of the test limits the amount of type II error allowed, which in turn creates an upper bound on the buyer's incentive to invest in equilibrium. Intuitively, if the buyer invests too often (which is correctly conjectured by the seller in equilibrium), then no signal can convince the seller to set the low price, which in turn kills all the buyer's ex-ante investment incentives. Therefore, information control cannot achieve the first-best.

In Section 4, we discuss how the findings above remain when the buyer's investment stochastically (rather than deterministically) affects his valuation. The optimal signal structure under stochastic investment continues to result in zero ex-post inefficiency because it is designed to never make any type II error. The only difference lies in the way type I errors generate ex-ante investment incentive for the buyer now, thus affecting the exact configuration (but not the form) of the optimal signal structure.

In Section 5, we allow the buyer to have more than one investment choice. We show that the separation between information that creates ex-ante investment incentive and information that causes ex-post inefficiency remains, so ex-post inefficiency is still always zero under the optimal signal structure. The buyer's ex-ante investment incentive is created when the seller is too pessimistic about the buyer's valuation and under-charges him, whereas ex-post inefficiency due to trade breaking down arises when the seller is too optimistic about the buyer's valuation and over-charges him. Pessimism is created when the information structure muddles information about the buyer's true investment with signals that suggest *lower* investments; on the other hand, optimism is created when information about the buyer's true investment is muddled with signals that suggest *higher* investments. Therefore, the optimal information structure is designed to allow for seller pessimism but eliminates seller optimism. Under the optimal signal structure, whenever the seller charges a price p, she is certain that the buyer's true valuation is at least p but is otherwise unsure about how high above p is the valuation.

The optimal signal structure is thus determined by the pessimism required to create the necessary investment incentive for the buyer. With only a single investment option, pessimism can only be created by generating the "low" signal some of the time when the buyer's valuation is high. Suppose there is an additional investment option that leads to a moderate valuation. Pessimism in a high-valuation state can now be created by generating the "moderate" signal some of the time, or a combination of both the "low" and "moderate" signals. This makes the characterization of the optimal signal structure significantly more difficult. To make progress, we restrict attention to a set of buyer investment strategy denoted by $\mathcal{Q}$. Intuitively, the restriction under $\mathcal{Q}$ puts an upper bound on the probability of taking a particular investment relative to the probability of taking the next lower investment. Under the single investment setting, $\mathcal{Q}$ corresponds to the set of all implementable investment strategies. We characterize the optimal signal structure that implements the investment strategies in $\mathcal{Q}$ and the set of social welfare achievable when there are more than one type of investment. Using an example with two investment options, we illustrate when and why the highest social welfare is achieved by implementing an investment strategy within $\mathcal{Q}$.

2 Related Literature

Our paper is primarily related to the literature on the use of asymmetric information to mitigate hold-up risks. Gibbons (1992) and Gul (2001) show that when the investment is completely unobservable to the seller, the ex-ante efficiency and ex-post inefficiency created from the asymmetric information exactly cancel out; thus the welfare is unchanged from the hold-up situation. Lau (2008) shows that the two effects change at different rates when the probability of the seller observing the investment outcome varies. Therefore, welfare can be improved if the seller observes the investment outcome with an intermediate probability. In addition, both Gul (2001) and Lau (2008) emphasize how welfare is always improved if the seller can make repeated and frequent offers after rejection by the buyer.

Our setting contrasts these two papers in two main ways. First, we allow for more general forms of information control and also illustrate that choosing or randomizing between perfect observability and perfect unobservability is never optimal. Second, we show that conditional on the optimal information structure, allowing for repeated offers upon rejection by the buyer has no effect since the seller's offer will be accepted immediately. Yet our results also complement these earlier works in that it illustrates why allowing for repeated offers in their setup *always* improves welfare – repeated offers eliminate ex-post inefficiency due to trade breakdown which occurs only in the "no-investment states", and our results indirectly point out that allowing for renegotiation at these states has no detrimental effect on the buyer's ex-ante investment incentive.

Other papers that study asymmetric information in the hold-up problem include González (2004), Hermalin and Katz (2009), Hermalin (2013), Halac (2015), and Tan (2017). As in Gul (2001), these papers study a variety of related issues while restricting attention to perfect observability versus perfect unobservability of the investment, but they do not consider more general forms of information control as in here.

Our paper is also related to the information control literature – see for example Rayo and Segal (2010), Ostrovsky and Schwarz (2010), Kamenica and Gentzkow (2011), and subsequent works on Bayesian persuasion. The difference is that these papers put no restriction on the signal structure choice (at least within the class of signal structures considered), whereas the signal structure in our setup has to also satisfy an equilibrium condition because information control is embedded in a hold-up problem. Consequently, we cannot appeal to the "concavification" argument (Aumann and Maschler, 1995), which is commonly used in the Bayesian persuasion settings; and (as will be discussed) the underlying logic of the optimal signal structure here will also be very different.

Away from pure information control, Condorelli and Szentes (2017) study information design in bilateral trade by allowing the buyer to publicly choose the distribution of his valuation and consider how this choice affects the buyer's ex-ante expected information rent; by contrast, we study the effects of the information transmission of the buyer's realized valuation instead. Bergemann, Brooks, and Morris (2015) study the effects of the signal structure that generates signals about the buyer's valuation to the seller; and Roesler and Szentes (2017) study the effects of the signal structure that generates signals about the buyer's valuation to the buyer himself. Both of these papers do not allow the buyer to determine his own valuation which is the focus of our paper.

3 Single Investment

Single type of investment is a special case of Section 5. However, we provide all the details in this section independent of Section 5, as this special case allows for a direct exposition while illustrating the main intuitions transparently. The omitted proofs for this section are in Appendix A.

3.1 Model

A buyer (he) has valuation $v = L$ for a good that a seller (she) can produce at a cost which is normalized to zero. Before interacting with the seller, the buyer can privately increase his valuation to $v = H$ at a cost c. Increasing the valuation is henceforth termed as an investment. We assume that $H - L > c$ so that it is socially efficient to invest. But due to incomplete contracts, the investment is not contractible.

After the investment decision is made but before trade, the seller receives a signal s regarding v. Let the set of signals be S. For expositional clarity, we assume that S is a finite set, although this is without loss of generality. A signal structure is defined by $\{S, \pi\}$ where $\pi\left(s|v\right)$ denotes the conditional probability of $s \in S$ given valuation v. This signal structure is common knowledge to both players at the start of the game.

After observing the signal, the seller makes a take-it-or-leave-it offer p to the buyer. If accepted, the seller's payoff is p while the buyer's payoff is $v - p - \mathbb{I}c$, where $\mathbb{I}$ is an indicator function that takes the value 1 if the buyer invested, and is zero otherwise; if rejected, the seller's payoff is 0 and the buyer's payoff is $-\mathbb{I}c$.

Our equilibrium concept is the Perfect Bayesian equilibrium. Given a signal structure, the buyer optimally chooses to invest or not, taking into account the distribution of signals that his investment decision will generate and his conjecture about the seller's pricing strategy after observing each signal $s \in S$. The seller, upon observing a signal s, forms a posterior which depends on both her conjecture about the buyer's investment strategy and the distribution of the signals under the signal structure, and then optimally sets a price based on her posterior. The buyer then accepts if $p \leq v$ and rejects otherwise. In equilibrium, each player's conjecture about the other player's strategy is correct. It is readily noted that the seller will only set $p = L$ or $p = H$ in equilibrium.

Throughout, we let $q \in [0, 1]$ denote the probability of the buyer investing. Thus, we say that *a signal structure implements* q if the buyer's strategy of investing with probability q can be sustained as an equilibrium under the signal structure. In addition, all payoffs will be expected payoffs; henceforth, we drop the "expected" quantifier to ease exposition.

Proposition 1. *The buyer's payoff is always zero in equilibrium. Moreover,* $q = 1$ *cannot be implemented.*

Proof. Since the seller will never set p lower than L, the buyer's payoff is zero if he does not invest (i.e. $q = 0$). If $q \in (0, 1)$ in equilibrium, the buyer must be indifferent between investing and not investing, which means that his payoff is also 0. Lastly, q cannot be 1 in equilibrium. This is because if $q = 1$, the seller will correctly conjecture it in equilibrium and always sets $p = H$, in which case the buyer will never invest in the first place. □

Proposition 1 thus implies that the seller's payoff in equilibrium is also the social welfare, so there is no need to differentiate between the two. Henceforth, the term "optimality" will refer to optimality of the seller's payoff. Since q cannot be 1 in equilibrium, the first-best optimality cannot be achieved.

3.2 Benchmarks

3.2.1 Fully Informative π: the Hold-up Case.

Consider a fully informative signal structure first, which gives the classic hold-up problem: since the seller perfectly knows v, she always sets $p = v$ and hence, the buyer never invests in equilibrium. The social welfare is thus L which is all given to the seller.

3.2.2 Fully Uninformative π.

Consider a fully uninformative signal structure next: $\pi(s|v) = \pi(s'|v)\ \forall v \in \{L, H\},\ s, s' \in S$. The following restates the result of Gibbons (1992) and Gul (2001) that the players' payoffs under no information are the same as in the hold-up case:[1]

Proposition 2. *The equilibrium under the fully uninformative signal structure is unique: the buyer invests with probability $\frac{L}{H}$ and the seller sets $p = L$ with probability $\frac{c}{H-L}$ after every (uninformative) signal. The seller's equilibrium payoff is L. Therefore, the players' payoffs are the same as in the hold-up case in expectation.*

Proof. From Proposition 1, $q \neq 1$. Similarly, $q \neq 0$; if $q = 0$, the seller will also correctly conjecture that in equilibrium and always sets $p = L$, in which case the buyer will deviate to choosing $q = 1$ instead. Let ρ be the probability that the seller sets price $p = L$ (ρ is independent of the signal since the signal has no information). Since $q \in (0, 1)$, the buyer must be indifferent between investing and not investing, which implies that $\rho(H - L) - c = 0 \iff \rho = \frac{c}{H-L} \in (0, 1)$. Thus the seller is also randomizing over H and L in equilibrium, which implies that she must be indifferent between the two prices: $L = qH \iff q = \frac{L}{H}$. $\square$

The seller's ignorance about the buyer's investment limits her ability to expropriate the gains from investment, thus improving ex-ante efficiency by (partially) restoring the buyer's ex-ante investment incentive. As a result, the buyer invests with positive probability. However, the asymmetric information at the trading stage creates ex-post inefficiency because trade breaks down when the buyer did not invest but the seller sets $p = H$. These two effects exactly cancel each other out in equilibrium.

3.3 Optimal Signal Structure

We first fix a $q \in (0, 1)$ and solve for the signal structure that gives the highest seller payoff while implementing q, assuming that such a signal structure exists. The subsequent variables will be dependent on q, but we omit the argument throughout to ease notation. Let β_s be the seller's posterior belief that $v = H$ after observing signal s under signal structure $\{S, \pi\}$. With q correctly conjectured by the seller in equilibrium,

$$\beta_s = \Pr(v = H|s) = \frac{\pi(s|H)\, q}{\pi(s|H)\, q + \pi(s|L)(1 - q)}. \tag{3.1}$$

[1] See problem 2.23 in Gibbons (1992), and Proposition 1 in Gul (2001).

Conditional on belief β, the seller's payoff from setting $p = H$ is βH, and that from $p = L$ is L. Denote $x_s := q\pi(s|H) + (1-q)\pi(s|L)$ as the ex-ante probability of signal s realizing. We say that a signal structure $\{S, \pi\}$ is *almost direct* if $S = \{l, n, h\}$ and π has the following properties:[2]

$$\begin{aligned} \text{if } x_l > 0, \text{ then } \beta_l < \tfrac{L}{H} \ ; \\ \text{if } x_n > 0, \text{ then } \beta_n = \tfrac{L}{H} \ ; \\ \text{if } x_h > 0, \text{ then } \beta_h > \tfrac{L}{H} \ . \end{aligned} \tag{3.2}$$

An almost direct signal structure produces an incentive compatible pricing recommendation for the seller almost all the time. The seller chooses $p = L$ when she observes $s = l$, chooses $p = H$ when she observes $s = H$, but she is indifferent between either price when she receives the neutral signal n.

Lemma 1. *Suppose a signal structure $\{\pi, S\}$ implements q and the seller's payoff in the equilibrium is V. There exists an almost direct signal structure that also implements q and the seller's payoff in the equilibrium is also V.*

Lemma 1 implies that it is without loss of generality to restrict attention to almost direct signal structures, which we will henceforth do so. The intuition behind is similar to the revelation principle. For any signal structure, the seller sets the same price whenever her posterior is less (resp. more) than $\frac{L}{H}$, so all signals that generate posteriors that are less (resp. more) than $\frac{L}{H}$ can be grouped together accordingly.[3]

We follow Kamenica and Gentzkow (2011) and frame the problem as choosing a distribution of posteriors $\{x_s, \beta_s\}_{\sum_s x_s = 1}$ that has to satisfy the Bayes plausibility constraint:

$$\sum_{s \in \{l,n,h\}} x_s \beta_s = q. \tag{3.3}$$

The original signal structure can then be backed out via $\pi(s|H) = \frac{\beta_s x_s}{q}$ and $\pi(s|L) = \frac{(1-\beta_s)x_s}{1-q}$. In addition, the signal structure must satisfy the equilibrium condition:

[2] Note that β_s is undefined when $x_s = 0$. For completeness, we specify the convention that $\beta_l = 0$ when $x_l = 0$, $\beta_n = \frac{L}{H}$ when $x_n = 0$, and $\beta_h = 1$ when $x_h = 0$. These beliefs form the seller's off-equilibrium beliefs when there is a detectable deviation from q by the buyer.

[3] Unlike the Bayesian persuasion literature, restricting the signal space to be the state space here is not without loss (at least at this stage). In the persuasion literature, the papers typically assume that the Receiver has a unique optimal action under each belief, or that the Receiver always takes the Sender-preferred action when the Receiver is indifferent. In this paper, however, it is not *a priori* clear what is a "Sender-preferred action" at a belief whereby the seller is indifferent between setting either price. This is because the seller's randomization strategy can disrupt the equilibrium and thus have implication on her ex-ante expected payoff.

$$\Big[\pi\left(l|H\right)+\sigma\pi\left(n|H\right)\Big]\left(H-L\right)-c=0, \tag{3.4}$$

whereby $\sigma\in[0,1]$ is the probability that the seller sets $p=L$ after observing $s=n$. The left hand side of (3.4) is the buyer's payoff from investing. For him to be indifferent between investing and not investing (so that $q\in(0,1)$ in equilibrium), this payoff must be the same as his payoff from not investing, which is 0. Using (3.1), condition (3.4) is equivalent to:

$$x_l\beta_l+\sigma x_n\beta_n=q\left(\frac{c}{H-L}\right). \tag{3.5}$$

Therefore, a signal structure is an almost direct signal structure that implements q and σ if the resulting distribution of posteriors $\{x_s,\beta_s\}_{\sum_s x_s=1}$ satisfies (3.2), (3.3) and (3.5). The resulting seller payoff is:

$$\left(x_l+x_n\right)L+x_h\beta_hH \;=\; L+x_h\left(\beta_hH-L\right) \tag{3.6}$$

Lemma 2. *Suppose there exists an almost direct signal structure with $x_n>0$ that implements q and $\sigma<1$, and the seller's payoff in the equilibrium is V. There exists an almost direct signal structure that implements q and $\sigma=1$, and the seller's payoff in the equilibrium is strictly higher than V.*

Proof. Suppose a distribution of posteriors $\{x_s,\beta_s\}$ supports an equilibrium with q and $\sigma<1$. Consider another distribution of posteriors $\{x'_s,\beta'_s\}$ whereby $x'_l=x_l+(1-\sigma)\,x_n\left(1-\frac{\beta_n}{\beta_h}\right)$, $x'_n=\sigma x_n$, $x'_h=x_h+(1-\sigma)\,x_n\frac{\beta_n}{\beta_h}$; $\beta'_l=\frac{x_l}{x'_l}\beta_l$, $\beta'_n=\beta_n$, $\beta'_h=\beta_h$. It is readily verified that $\{x'_s,\beta'_s\}$ satisfies (3.2) (3.3) and (3.5). Since $x'_h>x_h$ and $\beta'_h=\beta_h$, from (3.6), the seller's payoff under $\{x'_s,\beta'_s\}$ is higher. □

Intuitively, at the interim stage (after the signal realization but before trade), the seller's (interim) payoff under each signal can be ranked: signal l gives payoff L; signal n gives payoff L regardless of the seller setting $p=L$ or $p=H$; and signal h gives a payoff that is higher than L. Thus, if there is a probability of $1-\sigma>0$ that she will set $p=H$ after seeing signal n, this probability can be appropriately shifted to increase the ex-ante likelihood of signal h being realized which gives the seller a higher payoff. Doing so would not affect the equilibrium condition (3.5) since the probability of the seller offering $p=L$ after the buyer has invested will not be altered.

Lemma 2 implies that when searching for the (almost direct) signal structure that maximizes the seller's payoff, we can restrict attention to equilibria with $\sigma=1$. In turn, signals l

and n are essentially equivalent and can hence be pooled together. Thus we can restrict attention to *direct* signal structures whereby $S = \{l, h\}$, the resulting posteriors satisfy $\beta_l \leq \frac{L}{H}$ and $\beta_h > \frac{L}{H}$, and the seller plays $p = L$ at $s = l$ and $p = H$ at $s = h$.[4] The following theorem gives the main result of this section:

Theorem 1. *A signal structure that implements q exists if and only if $q \leq \frac{L}{L+c}$. For any $q \leq \frac{L}{L+c}$, the signal structure that maximizes the seller's payoff while implementing q is unique within the set of direct signal structures. It consists of:*

$$\begin{pmatrix} \pi(l|L) & \pi(h|L) \\ \pi(l|H) & \pi(h|H) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{c}{H-L} & 1 - \frac{c}{H-L} \end{pmatrix}$$

The resulting posteriors are $\beta_h = 1$ and $\beta_l = \frac{1}{1+\frac{1-q}{q}\left(\frac{H-L}{c}\right)} \leq \frac{L}{H}$. Trade always takes place, and the seller's payoff is $L + q(H - L - c)$.

We emphasize that the equilibrium existence condition in Theorem 1 takes into account all possible signal structures, not just direct signal structures. This thus provides an upper bound on the possible investment frequency in equilibrium. Moreover, the optimal signal structure is independent of q (although the posteriors generated and the resulting payoffs are dependent on q). This means that the direct signal structure that optimally implements q is the same for all implementable q.

Although this signal structure is reminiscent of the optimal signal structure in the leading "prosecutor" example in Kamenica and Gentzkow (2011) (hereafter KG) – in the sense that one state is always revealed (the "L" state here and the "guilty" state in KG) – the reasonings behind the two signal structures are very different. In particular, since the seller's payoff is convex in her belief here,[5] the optimal signal structure in the KG world would have been the fully informative one, but this would bring back the hold-up problem here. More generally, in Bayesian persuasion, when the fully informative signal structure is not optimal but there is scope for persuasion, the optimal signal structure optimally pools "favorable" states with "unfavorable" states while maintaining the credibility of the signals. In contrast, the optimal signal structure here determines the conditional probabilities at each state *separately*.

More precisely, the buyer's ex-ante investment incentive is provided via a probability of

[4] To see this formally, suppose an almost direct signal structure $\{x_s, \beta_s\}$ implements $q = 1$ and $\sigma = 1$. Consider another almost direct signal structure $\{x_s', \beta_s'\}$ such that $x_l' = x_l + x_n$, $\beta_l' = \frac{x_l\beta_l + x_n\beta_n}{x_l + x_n}$; $x_n' = 0$, $\beta_n' = \beta_n$; and $x_h' = x_h$ and $\beta_h' = \beta_h$. It is readily verified that $\{x_s', \beta_s'\}$ also satisfies (3.2), (3.3) and (3.5) and gives the same payoff to the seller as $\{x_s', \beta_s'\}$. Since $x_n' = 0$, signal n is irrelevant.

[5] At belief $\beta \leq \frac{L}{H}$, the seller's payoff is L; at belief $\beta \geq \frac{L}{H}$, her payoff is βH.

$\frac{c}{H-L}$ that his investment is not detected by the seller, in which case, he gets to keep the investment gains. This probability is set so that he is ex-ante indifferent between investing or not, which effectively pins down $\pi(\cdot|H)$. As for $\pi(\cdot|L)$, since $\pi(h|L)$ is the conditional probability of having ex-post inefficiency due to trade not taking place when the buyer did not invest and $\pi(\cdot|L)$ also does not affect the buyer's ex-ante investment incentive, $\pi(h|L)$ is set to zero to eliminate all ex-post inefficiency. The result that ex-ante investment incentive and ex-post inefficiency are taken care of separately by $\pi(\cdot|H)$ and $\pi(\cdot|L)$ respectively thus implies that there is no tradeoff between increasing ex-ante investment incentive and eliminating ex-post inefficiency.

Corollary 1. *For any implementable q, there is zero ex-post inefficiency under the optimal signal structure that implements q. The set of seller payoff that is achievable in equilibrium is* $\left[L, \frac{H}{c+L}L\right]$.

Given a fixed q, a signal structure here can be also viewed as a hypothesis test for the investment. The "false positive" type I error of the test is $(1-q)\pi(h|L)$, which is the ex-ante probability of detecting an investment when there is not; while the "false negative" type II error is $q\pi(l|H)$, which is the ex-ante probability of failing to detect the investment when there is one. When the test makes a type I error, the seller sets the high price which will be rejected by the buyer whose valuation remains low; thus ex-post efficiency arises. On the other hand, when the test makes a type II error, the seller sets the low price which leaves some surplus for the buyer whose valuation is high; the potential for this then feeds back as the buyer's ex-ante investment incentive. Therefore, under the optimal signal structure, the test never makes any type I error but allows for some type II error. However, to maintain the accuracy of the test to the seller, there cannot be too much type II error, which in turn creates an upper bound on the implementable q. Intuitively, when q increases, the need to maintain the credibility of signal l implies that signal l needs to detect state L more accurately and makes less mistake via wrongly detecting state H. However, this mistake (i.e. the type II error) is what creates ex-ante investment incentive for the buyer, so the need to improve the accuracy of l due to a higher q in turn destroys the buyer's ex-ante incentive to invest at all.

Finally, the simplicity of the optimal signal structure implies that it can be replicated by practical arrangements in many hold-up situations. The requirement is simply a technology that imperfectly searches for hard evidence of the presence of the investment. For example, in a vertical relationship in which the upstream supplier can make a cost-saving investment, this "technology" can be the random inspection of the supplier's facilities or the delegation of this inspection to a monitor who shirks occasionally.

4 Stochastic Investment

We have thus far only considered investment that is deterministic. In this section, we show that our results readily extend to stochastic investment. Suppose that instead of a binary investment decision, the buyer now gets to choose an investment level $\rho \in [0,1]$ at a cost $\phi(\rho)$, whereby ρ is the probability that his valuation increases from L to H. We make standard assumptions on the cost function: $\phi(\cdot)$ is strictly increasing and convex, with $\phi'(0) = 0$ and $\lim_{q\to 1} \phi'(q) = \infty$.

Let $f(\rho) := \rho[L + \phi'(\rho)] - L$; it is readily verified that there exists unique ρ^* such that $f(\rho^*) = 0$.

Proposition 3. *When investment is stochastic, a signal structure that implements investment ρ exists if and only if $\rho \le \rho^*$. For any $\rho \le \rho^*$, the signal structure that maximizes the social welfare while implementing ρ is unique within the set of direct signal structures. It consists of:*

$$\begin{pmatrix} \pi(l|L) & \pi(h|L) \\ \pi(l|H) & \pi(h|H) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\phi'(\rho)}{H-L} & 1 - \frac{\phi'(\rho)}{H-L} \end{pmatrix}$$

The resulting posteriors are $\beta_h = 1$ and $\beta_l = \frac{1}{1+\frac{1-\rho}{\rho}\left(\frac{H-L}{\phi'(\rho)}\right)} \le \frac{L}{H}$. Trade always takes place, and the set of social welfare that is achievable in equilibrium is $[L, L + \rho^*(H-L) - \phi(\rho^*)]$.

We only provide the arguments behind Proposition 3; the formal proof is omitted since it is mostly retracing the steps to arrive at Theorem 1, with a minor modification which we explain now.

Given an almost direct signal structure, the buyer's payoff from investment ρ is $\rho\Big[\pi(l|H) + \sigma\pi(n|H)\Big](H-L) - \phi(\rho)$. His optimal investment level is thus determined by the first order condition:

$$\Big[\pi(l|H) + \sigma\pi(n|H)\Big](H-L) - \phi'(\rho) = 0. \tag{4.1}$$

This first order condition (4.1) then replaces the equilibrium condition in (3.4). By replacing "q" with "ρ" and "c" with "$\phi'(\rho)$", the analysis in Section 3.3 follows through with a few cautions and qualifications about the interpretations.

First, the analysis in Section 3.3 only holds under the assumption that $c < H-L$, hence ρ must satisfy $\phi'(\rho) < H-L$; but this is not a problem because $\pi(l|H) + \sigma\pi(n|H) < 1$ (if not, the seller will never set $p = H$), so condition (4.1) implies that any implementable ρ must satisfy $\phi'(\rho) < H - L$. Second, the buyer's payoff is no longer always zero under stochastic investment, so the objective in (3.6) is now the social welfare rather than just the seller's payoff. Third, the upper bound on the implementable investment ρ^* is "analogous" to the upper bound on the probability of investment $q = \frac{L}{L+c}$ in Theorem 1; it is readily verified that $\rho^* = \frac{L}{L+\phi'(\rho^*)}$, and $\rho \le \frac{L}{L+\phi'(\rho^*)}$ if and only if $\rho \le \rho^*$.[6] Fourth, the social welfare under an implementable investment level ρ is $L + \rho(H - L) - \phi(\rho)$, and it is strictly increasing in ρ because, as observed above, any implementable ρ must satisfy $H - L > \phi'(\rho)$; this thus gives an analogous set of implementable social welfare (c.f. Corollary 1).

We note that if the signals reveal information on ρ rather than on the outcome,[7] then even the fully informative signal structure will result in some ex-post inefficiency. Thus, the elimination of ex-post inefficiency hinges on the assumption that the information structure generates signals about the investment outcome.

5 Multiple Types of Investment

This section considers the hold-up setting with multiple types of investment. The omitted proofs and details for this section are in Appendix B.

5.1 Model

Returning to deterministic investment as in Section 3, we suppose now that the buyer has $m + 1$ possible investment actions. Investment action $i \in M := \{0, 1, \ldots, m\}$ results in a valuation of v^i and incurs a cost of c^i for the buyer. Without loss of generality, we assume that $v^0 < v^1 < \cdots < v^m$; $c^0 = 0$; and $c^i > 0$ for any $i \ge 1$. Thus $i = 0$ corresponds to no investment, and $m = 1$ is the case considered in Section 3. Note that at this point, we do not require c^i to be increasing in i nor that m is the first best investment action.

As previously, an ex-ante determined signal structure generates signals about the buyer's valuation before the seller makes her price offer. The signal structure consists of $\{S, \pi\}$ whereby S is the signal space and $\pi(s|v^i)$ is the conditional probability of signal $s \in S$ after the buyer has chosen investment action i. The buyer's investment strategy now is a probability vector $\vec{q} = (q^0, q^1, \ldots, q^m)$ whereby q^i is the probability that the buyer chooses investment action i. His strategy is mixed if $q^i < 1\ \forall i$. In equilibrium, $\vec{q}$ is correctly

[6] This comes from noting that $\rho \le \frac{L}{L+\phi'(\rho)}$ if and only if $f(\rho) \le 0$, and f is strictly increasing.

[7] This is the setting consider in Hermalin and Katz (2009).

conjectured by the seller. It is readily noted that the seller will choose a price from only the set of possible ex-post valuations $\{v^0, v^1, \ldots, v^m\}$, and she will never offer a price v^i if she believes that $v = v^i$ with zero probability.

The following two statements are the analogues of Propositions 1 and 2 when there are multiple types of investment:

Proposition 4.

1. *In any equilibrium, the buyer's investment strategy involves $q^0 > 0$, and his payoff is zero.*

2. *Under the fully uninformative signal structure, the seller's equilibrium payoff is v^0. Therefore the players' payoffs are the same as in the hold-up case in expectation.*

Proposition 4.1 states that the buyer chooses not to invest with strictly positive probability in equilibrium. This is because the seller will correctly conjecture the lowest investment that is played with positive probability by the buyer in equilibrium, hence she will never charge a price below that resulting valuation. This implies that the buyer will get a negative payoff unless the lowest investment played is of zero cost, which is only possible if it is the no-investment action. In turn, since the buyer must be indifferent among any action played with positive probability and the no-investment action gives him a zero payoff, the buyer's payoff in equilibrium is zero. As in the single investment setting, the seller's payoff is the social welfare, so there is no need to differentiate between the two.

If the buyer's investment is perfectly observed by the seller, the hold-up problem arises: the only equilibrium outcome is the buyer chooses not to invest and is charged the price v^0. The social welfare with hold-up is v^0 which is all extracted by the seller. Proposition 4.2 generalizes Proposition 2: even when there are multiple types of investment, the fully uninformative signal structure still does not improve welfare relative to the hold-up result as just discussed.

5.2 Optimal Signal Structure

We consider the optimal signal structure next. As in Section 3.3, we proceed by first fixing a buyer strategy $\vec{q}$ under the assumption that it is implementable, and then we consider the optimal signal structure that implements it. It remains without loss of generality to restrict attention to direct signal structures which we describe next. This claim requires a more general argument than the one used in Lemma 2, but the intuition is similar, so we relegate the details for this claim to Appendix B.

The seller's posterior belief after observing signal s is a probability vector $\vec{\beta}_s = \{\beta_s^0, \beta_s^1, \ldots, \beta_s^m\}$, whereby β_s^i is the probability that the seller assigns to $v = v^i$, and the updating formula is:

$$\beta_s^i = \Pr\left(v = v^i | s\right) = \frac{\pi\left(s|v^i\right) q^i}{\sum_{j=0}^m \pi\left(s|v^j\right) q^j}. \tag{5.1}$$

Let $x_s = \sum_{i=0}^m q^i \pi\left(s|v^i\right)$ be the ex-ante probability of signal s being realized.

A direct signal structure consists of a signal space $S = M$ and a set of conditional probabilities π which results in posteriors that satisfy the following condition:

$$\text{For any } s \in M\text{: if } x_s > 0\text{, then } \sum_{j \geq s} \beta_s^j v^s \geq \sum_{j \geq i} \beta_s^j v^i \ \forall i \neq s. \tag{5.2}$$

The term $\sum_{j \geq i} \beta_s^j$ in (5.2) is $\Pr\left[v \geq v^i | s\right]$, which is the seller's subjective probability upon observing signal s that the buyer will accept price v^i; thus $\sum_{j \geq i} \beta_s^j v^i$ is her interim payoff from offering price v^i after signal s. Condition (5.2) thus implies that the direct signal structure provides signals that give incentive compatible price offer recommendations to the seller – the seller sets $p = v^s$ upon receiving signal s.

Next, analogous to condition (3.5), a signal structure must satisfy an equilibrium condition. In particular, a direct signal structure $\{M, \pi\}$ implements $\vec{q}$ if:

$$\text{For any } i \text{ such that } q^i > 0\text{:} \ \sum_{j \leq i} \pi\left(j|v^i\right)\left(v^i - v^j\right) - c^i = 0. \tag{5.3}$$

Condition (5.3) is the buyer's incentive compatibility condition – his payoff from playing any investment action i in which $q^i > 0$ must be zero (Proposition 4.1).

Under a direct signal structure $\{M, \pi\}$ and buyer strategy $\vec{q}$, the seller's payoff is the ex-ante expected social welfare. Since $\sum_{j \leq i}^m \pi\left(j|v^i\right)$ is the conditional probability of trade after the buyer has chosen investment action i, the ex-ante expected welfare is:

$$\sum_{i=0}^m q^i \left[\sum_{j \leq i}^m \pi\left(j|v^i\right)\right]\left(v^i - c^i\right). \tag{5.4}$$

We say that a signal structure π *achieves* a seller payoff W if π can implement a buyer investment strategy $\vec{q}$ such that the seller's payoff under $\vec{q}$ and π is W.

Theorem 2.

1. *Suppose that $\vec{q}$ is implementable and the direct signal structure $\{M, \pi\}$ optimally implements $\vec{q}$. It holds that for all i such that $q^i > 0$, $\pi(j|v^i) = 0$ for all $j > i$; and the seller's payoff is $\sum_{i=0}^{m} q^i (v^i - c^i)$.*

2. *The set of achievable seller payoff is a closed interval $[v^o, W^*]$. Moreover, if a signal structure π^* can achieve W^*, then π^* can also achieve any $W \in [v^0, W^*]$.*

Corollary 2. *By Theorem 2.1, under the optimal signal structure, the posterior satisfies* $\Pr(v \leq v^s | s) = 1$ *for any $s \in M$ such that $x_s > 0$. Therefore trade always takes place.*

$$
\begin{array}{c c}
 & \begin{array}{ccccc} s=0 & s=1 & s=2 & \cdots & s=m \end{array} \\
\begin{array}{c} v^0 \\ v^1 \\ v^2 \\ \vdots \\ v^m \end{array} &
\begin{pmatrix}
\boldsymbol{q^0 \pi_{00}} & \boxed{q^0\pi_{01}} & \boxed{q^0\pi_{02}} & \cdots & \boxed{q^0\pi_{0m}} \\
\underline{q^1\pi_{10}} & \boldsymbol{q^1 \pi_{11}} & \boxed{q^0\pi_{12}} & \cdots & \boxed{q^0\pi_{1m}} \\
\underline{q^2\pi_{20}} & \underline{q^2\pi_{21}} & \boldsymbol{q^2 \pi_{22}} & \ddots & \boxed{q^0\pi_{2m}} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\underline{q^m\pi_{m0}} & \underline{q^m\pi_{m1}} & \underline{q^m\pi_{m2}} & \cdots & \boldsymbol{q^m \pi_{mm}}
\end{pmatrix}
\end{array}
$$

Figure 5.1: Unconditional probabilities of the three types of outcomes. Ex-ante investment incentive relies on the "underlined" outcomes; ex-post inefficiency arises in the "boxed" outcomes.

Theorem 2.1 provides a partial characterization of the optimal signal structure. To understand it, consider the unconditional probabilities of the outcomes. Denoting $\pi(j|v^i)$ by π_{ij} (to simplify notation), Figure 5.1 shows a probability matrix whereby entry $q^i \pi_{ij}$ is the unconditional probability of the "*outcome*" in which the buyer had taken investment action i and signal j is realized. The set of outcomes can be categorized into three types: the "boxed outcomes", the "underlined outcomes" and the "diagonal outcomes". The boxed outcomes are in the upper triangular section of the matrix whereby the probabilities of their occurrences are boxed; in these outcomes, the seller is overly optimistic about the buyer's valuation and hence charges too high a price which then results in trade breakdown and thus ex-post inefficiency. On the other hand, the underlined outcomes are in the lower triangular section of the matrix whereby the probabilities of their occurrences are underlined; in these outcomes, the seller is too pessimistic about the buyer's valuation and hence undercharges him which then allows the buyer to keep some of the investment gain. Finally, the diagonal outcomes are the diagonals of the matrix; in these outcomes, the seller has the exact judgement of the buyer's valuation and hence charges him his true valuation which results in the seller extracting all the investment gain.

Given $\vec{q}$, the seller strictly prefers only the diagonal outcomes to occur. However, the buyer's ex-ante investment incentive can only come from the underlined outcomes. In particular, along each row, the underlined entries in the matrix (and only these underlined entries) have to satisfy the buyer's incentive compatibility condition in (5.3). Therefore, underlined outcomes must occur with positive probability. The other concern is the seller's incentive compatibility condition in (5.2). This requires that along each column, the seller believes that the outcome is sufficiently likely to be the diagonal outcome. Notice that moving all the probabilities of the boxed outcomes (where ex-post inefficiency occurs) to the diagonal outcomes actually helps to satisfy (5.2) and does not affect (5.3). This illustrates the separability between eliminating ex-post inefficiency and creating ex-ante investment incentive for the buyer. Therefore, under the optimal signal structure, the boxed outcomes (hence ex-post inefficiency) never occurs.

The characterization of the optimal signal structure is then completed by appropriately choosing the occurrences of the underlined outcomes to generate the buyer's investment incentive to implement an investment strategy $\vec{q}$. Given how these choices must interact with the occurrences of the diagonal outcomes to simultaneously satisfy both conditions (5.2) and (5.3), this problem is significantly more difficult when there are multiple investments, and there is not any general property that one can exploit.

To make progress, we henceforth assume that $v^i - v^{i-1} \geq c^i$ for all $i \geq 1$, and we restrict attention to the set of buyer investment strategy $\mathcal{Q}$, defined as the following:

Definition 1. $\mathcal{Q}$ is the set of buyer investment strategy $\vec{q}$ that satisfies:

$$q^0 \geq \frac{c^1}{v^0} q^1 \quad \text{and} \quad q^i \geq \frac{\left(v^i - v^{i-1}\right) c^{i+1}}{\left(v^i - v^{i-1} - c^i\right) v^i} q^{i+1} \quad \forall i \in \{1, 2, \ldots, m-1\}. \tag{5.5}$$

$\mathcal{Q}$ is a set of investment strategy whereby the probability of taking any investment action is bounded above relative to the probability of the next lower investment action. Recall that the seller's incentive compatibility condition in (5.2) requires that along each column in Figure 5.1, the diagonal outcome is sufficiently likely to occur relative to the underlined outcomes. Intuitively, the investment strategies in $\mathcal{Q}$ would satisfy (5.2) more "easily". Using the results in Theorem 1, it is readily verified that $\mathcal{Q}$ is the full set of implementable buyer investment strategy when $m = 1$.

We say that a seller payoff W is *$\mathcal{Q}$-achieveable* if there exists a signal structure π and a buyer investment strategy $\vec{q} \in \mathcal{Q}$ such that π implements $\vec{q}$ and the resulting seller payoff is W.

Proposition 5. *Every $\vec{q} \in \mathcal{Q}$ is optimally implementable by the following signal structure:*

$$\begin{gathered} \pi\left(0|v^0\right) = 1 \quad ; \quad \pi\left(j|v^0\right) = 0 \ \forall j > 0; \\ \forall i \geq 1 \text{:} \ \ \pi\left(i-1|v^i\right) = \tfrac{c^i}{v^i - v^{i-1}} \quad ; \quad \pi\left(i|v^i\right) = 1 - \tfrac{c^i}{v^i - v^{i-1}} \quad ; \quad \pi\left(j|v^i\right) = 0 \ \ \forall j \neq i-1, i. \end{gathered} \tag{5.6}$$

The set of $\mathcal{Q}$-achieveable seller payoff is $\left[v^0, \sum_{i=0}^{m} \frac{\alpha^i}{\sum_{j=0}^{m} \alpha^j}\left(v^i - c^i\right)\right]$, *where* $\alpha^0 = 1$ *and* $\alpha^i = \frac{v^0}{c^1}\left(\prod_{j=1}^{i-1} \frac{\left(v^j - v^{j-1} - c^j\right) v^j}{\left(v^j - v^{j-1}\right) c^{j+1}}\right)$ *for* $i \geq 1$.

The signal structure in (5.6) has the simple feature that the seller's pessimism after an investment action i is created via only believing that it is the next lower investment action $i-1$. In terms of the probability matrix in Figure 5.1, under this information structure, only two outcomes arise with positive probability along each row $i \geq 1$: the diagonal outcome and the underlined outcome immediately to the left of it. Therefore, we call the signal structure in (5.6) an *adjacent type-II error signal structure* (hereafter *A2 signal structure*). In the proof of Proposition 5, we also show that any strategy that is implementable by the A2 signal structure must be in the set $\mathcal{Q}$.

As higher investment actions are also more efficient by assumption here,[8] the optimal investment strategy (i.e. the one that maximizes the seller's payoff) is the one whereby all the bounds in condition (5.5) are binding. This corresponds to an investment strategy in which $q^i = \frac{\alpha^i}{\sum_{j=0}^{m} \alpha^j}$; hence α^i represents the "weight" for investment action i under the optimal investment strategy.

The question that naturally arises next is whether if there exists any $\vec{q} \notin \mathcal{Q}$ that is implementable and results in a higher payoff for the seller. While we deem this question impossible to answer in general, we can consider the example of $m = 2$ whereby we can analytically solve for the signal structure that achieves the highest seller payoff. We show that

[8] $v^i - v^{i-1} \geq c^i$ implies that $v^i - c^i > v^{i-1} - c^{i-1}$.

for certain parameter values, the highest seller payoff is indeed $\mathcal{Q}$-achieveable; in other words, our restriction to the set $\mathcal{Q}$ in such instances is without loss and Proposition 5 characterizes the optimal signal structure that implements any implementable buyer strategy. In cases where this highest seller payoff is not $\mathcal{Q}$-implementable, we use the example to illustrate why and how better investment strategies can be implemented.

Example: $m = 2$.

We characterize the signal structure that implements the highest seller payoff under $m = 2$ now. Note that Theorem 2 together with condition (5.3) completely pin down $\pi(\cdot|v^0)$ and $\pi(\cdot|v^1)$; and $\pi(\cdot|v^2)$ is determined by $\pi(0|v^2)(v^2 - v^0) + \pi(1|v^2)(v^2 - v^1) = c^2$. Denote $\pi(0|v^2)$ by $\gamma \in \left[0, \frac{c^2}{v^2 - v^0}\right]$, and let $\pi(1|v^2) = f(\gamma) = \frac{c^2 - \gamma\left(v^2 - v^0\right)}{v^2 - v^1}$. Thus, for any implementable $\vec{q}$, the signal structure that implements it is parametrized by γ (which is dependent on $\vec{q}$):

$$\begin{pmatrix} \pi(0|v^0) & \pi(1|v^0) & \pi(2|v^0) \\ \pi(0|v^1) & \pi(1|v^1) & \pi(2|v^1) \\ \pi(0|v^2) & \pi(1|v^2) & \pi(2|v^2) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{c^1}{v^1 - v^0} & \frac{v^1 - v^0 - c^1}{v^1 - v^0} & 0 \\ \gamma & f(\gamma) & 1 - \gamma - f(\gamma) \end{pmatrix} \tag{5.7}$$

Proposition 6. *Let* $\xi = \frac{c^2\left(v^1 - v^0\right) + v^0\left(v^2 - v^0\right)}{(c^1 + v^0)(v^2 - v^0) - (v^1 - v^0 - c^1)v^1}$. *The signal structure that achieves the highest seller payoff under $m = 2$ is uniquely the signal structure in (5.7) whereby:*[9]

- *if* $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} \leq \xi$, *then* $\gamma = 0$;
- *if* $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} > \xi$, *then* $\gamma = \frac{c^1 c^2}{(v^1 - v^0)(v^2 - v^1 + c^1)}$.

Therefore, when $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} \leq \xi$, *the highest achievable seller payoff is $\mathcal{Q}$-achieveable.*[10]

We provide a bit of details of the derivation to outline the intuitions of Proposition 6; the full derivation is in the proof. Fixing a γ, the highest achievable seller payoff is the solution

[9] To be precise, the uniqueness property does not apply at the knife-edge case of $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} = \xi$. In this case, γ is anything in the set $\left[0, \frac{c^1 c^2}{(v^1 - v^0)(v^2 - v^1 + c^1)}\right]$.

[10] This is because the signal structure under $\gamma = 0$ is the A2 signal structure. As mentioned above, any strategy that is implementable by the A2 signal structure must be in $\mathcal{Q}$.

to program $\mathcal{P}(\gamma)$:

$$W(\gamma) \quad := \quad \max_{q^1;q^2} \;\; (1-q^1-q^2)\,v^0 + q^1\,(v^1-c^1) + q^2\,(v^2-c^2)$$

subject to

$$\left[(1-q^1-q^2) + q^1\left(\tfrac{c^1}{v^1-v^0}\right) + q^2\gamma\right] v^0 \geq \left[q^1\left(\tfrac{c^1}{v^1-v^0}\right) + q^2\gamma\right] v^1 \tag{5.8}$$

$$\left[(1-q^1-q^2) + q^1\left(\tfrac{c^1}{v^1-v^0}\right) + q^2\gamma\right] v^0 \geq q^2\gamma v^2 \tag{5.9}$$

$$\left[q^1\left(\tfrac{v^1-v^0-c^1}{v^1-v^0}\right) + q^2 f(\gamma)\right] v^1 \geq q^2 f(\gamma)\, v^2 \tag{5.10}$$

$$q^1 + q^2 \leq 1 \tag{5.11}$$

The objective follows from (5.4). Constraints (5.8) and (5.9) are conditions (5.2) for $s=0$; and (5.10) is condition (5.2) for $s=1$. We shall illustrate the point by starting at $\gamma=0$ (i.e. the A2 signal structure) and see if the seller's payoff can be improved by increasing γ.

First note that we can write (5.8) as $\frac{(1-q^1-q^2)}{q^1\frac{c^1}{v^1-v^0}+q^2\gamma}+1 \geq \frac{v^1}{v^0}$, which is violated if $q^1+q^2>1$; thus (5.11) is subsumed by (5.8). Moreover, this is a linear program, so the solution lies on a vertex. It is readily seen that constraint (5.9) is trivially satisfied at $\gamma=0$. So the solution is the intersection of (5.8) and (5.10) when both constraints bind. We note that these two constraints maintain the credibility of signals $s=0$ and $s=1$ respectively.

Consider increasing γ slightly from 0 to ε. Since $f(\varepsilon) < f(0)$, this relaxes constraint (5.10) which allows us to increase q^2 slightly at the expense of q^1. On the other hand, increasing ε and q^2 will violate constraint (5.8); to restore it, we have to decrease q^1 and move the probability weight to q^0. Denote the increase in q^2 by z^2; and the decrease in q^1 by z^1. The resulting change in the seller's payoff is thus:

$$\left(z^1-z^2\right)v^0 - z^1\left(v^1-c^1\right) + z^2\left(v^2-c^2\right) = z^2\left(v^2-v^0-c^2\right) - z^1\left(v^1-v^0-c^1\right).$$

Therefore, increasing γ from 0 to ε will not increase the seller's payoff if:

$$\frac{v^2-v^0-c^2}{v^1-v^0-c^1} \leq \frac{z^2}{z^1}. \tag{5.12}$$

Condition (5.12) has the intuitive interpretation that the left-hand side is the ratio of the value of investment action 2 relative to 1; while the right-hand side is the ratio of the increase in likelihood of 2 relative to the decrease in likelihood of 1.

The changes in probabilities (i.e. values of z^1 and z^2) depend on ε. For small enough ε, it remains true that the solution to program $\mathcal{P}(\varepsilon)$ is the intersection of the two binding constraint (5.8) and (5.10). Thus, z^1 and z^2 are determined by these two conditions. Since (5.8) must bind for both $\gamma = 0$ and $\gamma = \varepsilon$, we have:

$$z^1 \left(c^1 + v^0\right) - z^2 \left(v^0 + \varepsilon \left[v^1 - v^0\right]\right) = \hat{q}^2 \varepsilon \left[v^1 - v^0\right], \tag{5.13}$$

where $\hat{q}^2$ is the solution of q^2 under $\gamma = 0$. Likewise, (5.10) must bind for both $\gamma = 0$ and $\gamma = \varepsilon$, which gives:

$$z^2 \left(c^2 - \varepsilon \left[v^2 - v^0\right]\right) \left(v^1 - v^0\right) + z^1 \left(v^1 - v^0 - c^1\right) v^1 = \hat{q}^2 \varepsilon \left(v^2 - v^0\right) \left(v^1 - v^0\right) \tag{5.14}$$

Combining (5.13) and (5.14) and taking ε to the zero limit, we get:

$$\frac{z^2}{z^1} = \frac{c^2 \left(v^1 - v^0\right) + v^0 \left(v^2 - v^0\right)}{\left(c^1 + v^0\right)\left(v^2 - v^0\right) - \left(v^1 - v^0 - c^1\right) v^1} = \xi. \tag{5.15}$$

Condition (5.15) provides the ratio in the changes in likelihood of investment actions 2 and 1 when increasing γ slightly from 0. Combining (5.15) with (5.12), we can conclude that increasing γ slightly from 0 worsens the seller's payoff when $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} \leq \xi$.

This tradeoff, which compares the effects of relative likelihood of investment actions 2 and 1 (i.e. $\frac{z^2}{z^1}$) with the relative payoffs from the two actions (i.e. $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1}$), determines whether the seller's payoff can be improved by increasing γ. For low γ in which constraints (5.8) and (5.10) are binding – which happens when $\gamma \leq \frac{c^1 c^2}{(v^1 - v^0)(v^2 - v^1 + c^1)}$ – the effect of a marginal increase in γ on the relative likelihood of investment actions (i.e. $\frac{z^2}{z^1}$) is always the expression in (5.15). Therefore, increasing γ increases (resp. decreases) the seller's payoff when $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} < \xi$ (resp. $\frac{v^2 - v^0 - c^2}{v^1 - v^0 - c^1} > \xi$).

When $\gamma > \frac{c^1 c^2}{(v^1 - v^0)(v^2 - v^1 + c^1)}$, constraint (5.9) becomes binding, so the effect of an increase in γ on the relative likelihood of investment actions changes, and we can verify that it is always lower than the relative payoffs. This implies that increasing γ beyond $\frac{c^1 c^2}{(v^1 - v^0)(v^2 - v^1 + c^1)}$ always decreases the seller's payoff.

6 Conclusion

The literature has noted that introducing information asymmetry regarding the buyer's investment can prevent the seller from abusing her bargaining power and hence alleviate the hold-up risks. Implicitly suggested in these earlier papers is a tradeoff between ex-ante investment incentive and ex-post inefficiency due to the asymmetric information. In this paper, we make the point that such a tradeoff is unnecessary because the information that creates ex-ante investment incentive (when hidden) is different from the information that creates ex-post inefficiency (when hidden). Consequently, by hiding and revealing the right information, ex-post inefficiency can be eliminated without compromising the ex-ante investment incentive. Moreover, such forms of more efficient information control often do not require overly complex arrangements in the economic relationship. In turn, we hope that our results can serve as a guidance for future work on how to better make use of information control to mitigate the hold-up problem in various applied settings.

A Appendix for Section 3

Proof of Lemma 1

Proof. Let $S_h \subset S$ be the set of signals such that the posteriors generated are strictly greater than $\frac{L}{H}$; analogously, let S_l (resp. S_n) be the set of signals with posteriors strictly less than (resp. equal to) $\frac{L}{H}$. In addition, let σ_s denote the probability of the seller playing $p = L$ after observing signal s. In equilibrium, σ_s must be 1 if $s \in S_l$, and it must be 0 if $s \in S_h$, while it can be anything between 0 to 1 when $s \in S_n$. Consider the following almost direct signal structure $\left\{S^{ad}, \pi^{ad}\right\}$ where $S^{ad} := \{l, n, h\}$: $\pi^{ad}\left(s^{ad}|v\right) = \sum_{s \in S_{s^{ad}}} \pi\left(s|v\right)$ for all $s^{ad} \in S^{ad}$; let the seller play $p = L$ with probability $\sigma^{ad} = \frac{\sum_{s \in S_n} \pi[s|H]\sigma_s}{\sum_{s \in S_n} \pi[s|H]} = \frac{\sum_{s \in S_n} \pi[s|H]\sigma_s}{\pi^{ad}[n|H]}$ upon observing n.[11]

To check that this is an equilibrium with the buyer investing with probability q, first note that given q, the seller's pricing strategy is clearly a best response. For the buyer, his expected payoff after investing is

$$\begin{aligned}
\Pr\left[p = L|v = H\right](H - L) - c &= \left(\pi^{ad}\left(l|H\right) + \pi^{ad}\left(n|H\right)\sigma^{ad}\right)(H - L) - c \\
&= \left(\sum_{s \in S_l} \pi\left[s|H\right] + \sum_{s \in S_n} \pi\left[s|H\right]\sigma_s\right)(H - L) - c,
\end{aligned}$$

where the second line is the buyer's expected payoff after investing under the original signal structure $\{S, \pi\}$. The buyer's payoff when he does not invest is 0 under both signal structures. Since the buyer is indifferent between investing or not under $\{S, \pi\}$, he is also indifferent under $\left\{S^{ad}, \pi^{ad}\right\}$. Thus q is the buyer's best response as well, and hence it is an equilibrium.

We check the payoff next. Conditional on the buyer investing, the seller's expected payoff under $\left\{S^{ad}, \pi^{ad}\right\}$ is:[12]

$$\left(\pi^{ad}\left(l|H\right) + \pi^{ad}\left(n|H\right)\right)L + \pi\left(h|H\right)H = \left(\sum_{s \in S_l} \pi\left(s|H\right) + \sum_{s \in S_n} \pi\left(s|H\right)\right)L + H \cdot \sum_{s \in S_h} \pi\left(s|H\right)$$

where the RHS is seller's expected payoff, conditional on the buyer investing, under $\{S, \pi\}$. Next, conditional on the buyer not investing, the seller's expected payoff under $\left\{S^{ad}, \pi^{ad}\right\}$

[11] If S_n is empty, then this is irrelevant.

[12] Note that the seller's conditional expected payoff under posterior $\frac{L}{H}$ is always L.

is:

$$\left(\pi^{ad}\left(l|L\right)+\pi^{ad}\left(n|L\right)\right)L+\pi\left(h|H\right)\cdot 0=\left(\sum_{s\in S_l}\pi\left(s|L\right)+\sum_{s\in S_n}\pi\left(s|L\right)\right)L+0\cdot\sum_{s\in S_h}\pi\left(s|H\right),$$

where RHS is seller's expected payoff, conditional on the buyer not investing, under $\{S,\pi\}$.

□

Proof of Theorem 1

Proof. We prove the "only if" direction of the existence result first. Suppose, for a contradiction, that $q>\frac{L}{L+c}$ but there exists an almost direct signal structure that implements q. From Lemma 2, there exists a direct signal structure that implements q. Let $\beta_l\leq\frac{L}{H}$ and $\beta_h>\frac{L}{H}$ be the resulting posteriors. From (3.5), $\beta_l=\frac{qc}{x_l(H-L)}$; from (3.3), $x_l=\frac{\beta_h-q}{\beta_h-\beta_l}$. Combining the two, we get:

$$\beta_l=\frac{q\left(\frac{c}{H-L}\right)\beta_h}{\beta_h-q+q\left(\frac{c}{H-L}\right)}=\frac{q\left(\frac{c}{H-L}\right)}{1-\frac{q}{\beta_h}\left(1-\frac{c}{H-L}\right)}\tag{A.1}$$

Since $\frac{c}{H-L}<1$, β_l is decreasing in β_h. $\beta_h\leq 1$ then implies that $\beta_l\geq\frac{q\left(\frac{c}{H-L}\right)}{1-q\left(1-\frac{c}{H-L}\right)}$. When $q>\frac{L}{L+c}$, $\beta_l>\frac{L}{H}$ which contradicts $\beta_l\leq\frac{L}{H}$. Next, for the "if" direction, it is readily verified that when $q\leq\frac{L}{L+c}$, the signal structure in the theorem results in posteriors $\beta_h=1$ and $\beta_l\leq\frac{L}{H}$, and it satisfies the equilibrium condition (3.5).

For optimality, it suffices to consider the set of direct signal structures (Lemma 2). For any β_h, the corresponding β_l is (A.1), and the seller's payoff, from (3.6), is $V\left(\beta_h\right)=L+x_h\beta_hH-x_hL$. From (3.3) and (3.5):

$$\begin{aligned}x_h\beta_hH &= \left(q-x_l\beta_l\right)H = \left[q-q\left(\tfrac{c}{H-L}\right)\right]H\\ x_hL &= \tfrac{q-\beta_l}{\beta_h-\beta_l}L = \left(\frac{q-\frac{q\left(\frac{c}{H-L}\right)\beta_h}{\beta_h-q+q\left(\frac{c}{H-L}\right)}}{\beta_l-\frac{q\left(\frac{c}{H-L}\right)\beta_h}{\beta_h-q+q\left(\frac{c}{H-L}\right)}}\right)L = \tfrac{q}{\beta_h}\left(1-\tfrac{c}{H-L}\right)L\end{aligned}$$

Therefore, $V\left(\beta_h\right)=L+q\left(H-\frac{L}{\beta_h}\right)\left[1-\left(\frac{c}{H-L}\right)\right]$. Since $V\left(\beta_h\right)$ is strictly increasing, the optimal β_h is 1, and $\beta_l=\frac{q\left(\frac{c}{H-L}\right)}{1-q+q\left(\frac{c}{H-L}\right)}$. The seller's payoff is $V\left(1\right)=L+q\left(H-L-c\right)$. The signal structure is then backed out via $\pi\left(s|H\right)=\frac{\beta_sx_s}{q}$ and $\pi\left(s|L\right)=\frac{(1-\beta_s)x_s}{1-q}$. □

B Appendix for Section 5

Proof of Proposition 4

Proof. Statement 1: First note that the buyer obtains a zero payoff by not investing (i.e. choosing $q^0 = 1$), so his equilibrium payoff is weakly higher than zero. Suppose for a contradiction that $q^i = 1$ for an $i \geq 1$. Since the seller correctly conjectures the buyer's strategy in equilibrium, she will set $p = v^i$ which means that the buyer's ex-ante payoff is $-c^i < 0$ (contradiction). Thus, in equilibrium, either $q^0 = 1$ or the buyer's strategy is mixed. Clearly the proposition is true if $q^0 = 1$. Suppose the buyer's strategy is mixed now. Let $\underline{i}$ be the lowest investment action played with strictly positive probability. In the mixed strategy equilibrium, the buyer is indifferent between playing any investment action i with $q^i > 0$; thus his payoff is that of choosing investment $\underline{i}$. Since the seller correctly conjectures $\vec{q}$, her posterior belief must satisfy $\Pr\left[v < v^{\underline{i}}\right] = 0$; thus she will never offer a price lower than $v^{\underline{i}}$, which in turn implies that the buyer's payoff is no higher than $-c^{\underline{i}}$. Since his payoff must be at least 0, $\underline{i}$ must be 0 which implies $q^0 > 0$. In turn, since the seller will never charge $p < v^0$, the buyer's payoff is zero.

Statement 2: If the seller offers $p = v^0$, her payoff is v^0 since trade is guaranteed. Thus it suffices to show that the seller's equilibrium strategy must involve offering $p = v^0$ with strictly positive probability. To show this, note that if the lowest price offered by the seller in equilibrium is some $v^{\underline{i}} > v^0$, the buyer must be choosing investment $\underline{i} > 0$ with strictly positive probability. But since the price is never below $v^{\underline{i}}$, the buyer's payoff from choosing investment $\underline{i}$ is $-c^{\underline{i}}$. Since the buyer's payoff in equilibrium is 0, $\underline{i}$ must be 0. □

Sufficiency of Direct Signal Structures

Fix a buyer investment strategy $\vec{q} = (q^0, q^1, \ldots, q^m)$ which is correctly conjectured by the seller in equilibrium. Under an arbitrary signal structure $\{S, \pi\}$, the seller's posterior upon observing signal $s \in S$ is a probability vector $\vec{\beta}_s = (\beta_s^0, \beta_s^1, \ldots, \beta_s^m)$ whereby β_s^i is the probability that the seller assigns to $v = v^i$, and the updating formula is in (5.1) in the main text. The ex-ante probability of signal s is thus $x_s = \sum_{i=0}^m q^i \pi\left(s|v^i\right)$. Let the seller's strategy be $\vec{\sigma} := \{\vec{\sigma}_s\}_{s \in S}$, whereby $\vec{\sigma}_s = (\sigma_s^0, \sigma_s^1, \ldots, \sigma_s^m)$ is a probability vector and σ_s^i is the probability that the seller offers price $p = v^i$ upon observing signal s.[13]

The signal structure has to satisfy equilibrium conditions whereby both players' strategies

[13] The strategy space of the seller depends on the signal space S but we omit the argument to ease notation.

are best responses against each other. $\vec{\sigma}$ is a best response price strategy for the seller if:

$$\text{For all } s \in S \text{, if } \sigma_s^i > 0 \text{, then } \sum_{k \geq i} \beta_s^k v^i \geq \sum_{k \geq j} \beta_s^k v^j \quad \forall j \neq i,$$

which can be equivalently written as:

$$\text{For all } s \in S \text{, if } \sigma_s^i > 0 \text{, then } \sum_{k \geq i} q^k \pi\left(s|v^k\right) v^i - \sum_{k \geq j} q^k \pi\left(s|v^k\right) v^j \geq 0 \quad \forall j \neq i. \tag{B.1}$$

For the buyer, for every investment played with strictly positive probability under $\vec{q}$, his expected payoff from it must be 0 (Proposition 4). This is equivalent to:

$$\text{For all } i \in M \text{, if } q^i > 0 \text{, then } \sum_{s \in S} \pi\left(s|v^i\right) \sum_{j \leq i} \sigma_s^j \left(v^i - v^j\right) = c^i \ . \tag{B.2}$$

Thus, a signal structure implements $\vec{q}$ and $\vec{\sigma}$ if the signal structure (and its resulting posteriors) satisfies (B.1) and (B.2). The seller's ex-ante expected payoff in equilibrium is:

$$\sum_{s \in S} x_s \left[\max_{i \in M} \sum_{k \geq i} \beta_s^k v^i \right].$$

Let Σ_S be the set of pure strategies of the seller under a signal space S – that is, if $\vec{\sigma} \in \Sigma_S$, then for all $s \in S$, there exists $i \in M$ such that $\sigma_s^i = 1$.

Lemma 3. *Suppose that the signal structure $\{S, \pi\}$ implements $\vec{q}$ and $\vec{\sigma} \notin \Sigma_S$, and it gives the seller a payoff of V. There exists a signal structure $\{M, \hat{\pi}\}$ that implements $\vec{q}$ and $\vec{\hat{\sigma}} \in \Sigma_M$, whereby $\hat{\sigma}_i^i = 1$ $\forall i \in M$; and it also gives the seller a payoff of V.*

Proof. For each $\hat{s} \in M$ and $i \in M$, set:

$$\hat{\pi}\left(\hat{s}|v^i\right) = \sum_{s \in S} \pi\left(s|v^i\right) \sigma_s^{\hat{s}}.$$

For any $i \in M$, $\sum_{\hat{s} \in M} \left[\sum_{s \in S} \pi\left(s|v^i\right) \sigma_s^{\hat{s}}\right] = \sum_{s \in S} \pi\left(s|v^i\right) = 1$, so $\hat{\pi}$ is a valid signal structure. $\vec{\hat{\sigma}}$ is the seller's best response if $\forall i \neq j$:

$$\begin{aligned}
& \textstyle\sum_{k \geq i} q^k \hat{\pi}\left(i|v^k\right) v^i \geq \sum_{k \geq j} q^k \hat{\pi}\left(i|v^k\right) v^j \\
\iff \quad & \textstyle\sum_{k \geq i} q^k \left[\sum_{s \in S} \pi\left(s|v^k\right) \sigma_s^i\right] v^i \geq \sum_{k \geq j} q^k \left[\sum_{s \in S} \pi\left(s|v^k\right) \sigma_s^i\right] v^j \\
\iff \quad & \textstyle\sum_{s \in S} \left[\sum_{k \geq i} q^k \pi\left(s|v^k\right) v^i - \sum_{k \geq j} q^k \pi\left(s|v^k\right) v^j\right] \sigma_s^i \geq 0
\end{aligned}$$

where the last inequality holds from (B.1). Next, if playing $\vec{q}$ is best response for the buyer under $\hat{\pi}$, then for all $i \in M$, if $q^i > 0$, $\hat{\pi}$ satisfies:

$$\begin{aligned} & \textstyle\sum_{j \leq i} \hat{\pi}\left(j|v^i\right)\left(v^i - v^j\right) = c^i \\ \iff \quad & \textstyle\sum_{j \leq i} \left[\sum_{s \in S} \pi\left(s|v^i\right) \sigma_s^j\right]\left(v^i - v^j\right) = c^i \end{aligned}$$

which holds from (B.2). Therefore, $\hat{\pi}$ also implements $\vec{q}$. To check that the seller's payoff is also V under $\hat{\pi}$, recall from Proposition 4 that the seller's payoff is the social welfare. Thus, it suffices to check that the probabilities of trade breaking down at each valuation are the same across π and $\hat{\pi}$. Under π, conditional on v^i, the probability of no trade is $\sum_{s \in S} \pi\left(s|v^i\right) \sum_{j \geq i} \sigma_s^j$; under $\hat{\pi}$, the corresponding probability is $\sum_{j \geq i} \hat{\pi}\left(j|v^i\right) = \sum_{j \geq i}\left[\sum_{s \in S} \pi\left(s|v^i\right) \sigma_s^j\right] = \sum_{s \in S} \pi\left(s|v^i\right) \sum_{j \geq i} \sigma_s^j$. □

This thus establishes that it is without loss to restrict attention to direct signal structures as is done in the main text.

Proof of Theorem 2

Proof. Suppose π implements $\vec{q}$; thus π satisfies (5.2) and (5.3). Consider the following signal structure $\{M, \hat{\pi}\}$:

$$\text{For any } i, j \in M, \;\; \hat{\pi}\left(j|v^i\right) = \begin{cases} \pi\left(j|v^i\right) & \text{, if } j < i \\ \sum_{j' \geq i}^m \pi\left(j'|v^i\right) & \text{, if } j = i \\ 0 & \text{, if } j > i \end{cases}$$

Under $\hat{\pi}$, the buyer's payoff from playing investment i is $\left[\sum_{j \leq i}^m \hat{\pi}\left(j|v^i\right)\left(v^i - v^j\right)\right] - c^i = \left[\sum_{j \leq i}^m \pi\left(j|v^i\right)\left(v^i - v^j\right)\right] - c^i$. Thus $\hat{\pi}$ satisfies (5.3). Next, we check that $\hat{\pi}$ satisfies (5.2), which can be equivalently written as:

$$\sum_{k \geq i}^m q^k \pi\left(i|v^k\right) v^i \geq \sum_{k \geq j}^m q^k \pi\left(i|v^k\right) v^j \quad \text{, for all } i, j \in M \tag{B.3}$$

Fix a $i \in M$. For any $j < i$, $\sum_{k \geq j}^m q^k \hat{\pi}\left(i|v^k\right) = \sum_{k \geq i}^m q^k \hat{\pi}\left(i|v^k\right)$ since $q^k \hat{\pi}\left(i|v^k\right) = 0$ for any $k < i$; so $\sum_{k \geq i}^m q^k \hat{\pi}\left(i|v^k\right) v^i \geq \sum_{k \geq j}^m q^k \hat{\pi}\left(i|v^k\right) v^j$. Consider $j > i$ next. Note that

$\hat{\pi}\left(i|v^k\right) = \pi\left(i|v^k\right)$ for any $k > i$, while $\hat{\pi}\left(i|v^i\right) \geq \pi\left(i|v^i\right)$; so

$$\begin{aligned}&\sum_{k\geq i}^{m} q^k \hat{\pi}\left(i|v^k\right) v^i - \sum_{k\geq j}^{m} q^k \hat{\pi}\left(i|v^k\right) v^j \\ &= \left[q^k \hat{\pi}\left(i|v^i\right) v^i - q^k \pi\left(i|v^i\right) v^i\right] + \underbrace{\sum_{k\geq i}^{m} q^k \pi\left(i|v^k\right) v^i - \sum_{k\geq j}^{m} q^k \pi\left(i|v^k\right) v^j}_{\geq 0 \text{ from (B.3)}} \quad \geq \ 0\end{aligned}$$

Thus $\sum_{k\geq i}^{m} q^k \hat{\pi}\left(i|v^k\right) v^i = \sum_{k\geq j}^{m} q^k \hat{\pi}\left(i|v^k\right) v^j \ \forall j \neq i$, which hence satisfies (5.2).

This means that $\hat{\pi}$ also implements $\vec{q}$, and the ex-ante payoff of the seller is:

$$\sum_{i=0}^{m} q^i \left[\sum_{j\leq i}^{m} \hat{\pi}\left(j|v^i\right)\right]\left(v^i - c^i\right) = \sum_{i=0}^{m} q^i \left(v^i - c^i\right) \geq \sum_{i=0}^{m} q^i \left[\sum_{j\leq i}^{m} \pi\left(j|v^i\right)\right]\left(v^i - c^i\right);$$

whereby the inequality is strict if there exists $j > i$ such that $q^i > 0$ but $\pi\left(j|v^i\right) > 0$ (i.e. $\sum_{j\leq i}^{m} \pi\left(j|v^i\right) < 1$).

The highest achievable seller payoff is the solution to the program that chooses π and $\vec{q}$ to maximize (5.4) subject to constraints (5.2) and (5.3). The feasible set is clearly compact and the objective function is continuous; thus the maximum exists. Let the maximum payoff be W^* and suppose it is obtained under buyer strategy $\vec{q}^*$ which is implementable by π^*. Let $\bar{i}$ be the highest investment action that is played with strictly positive probability under $\vec{q}^*$. It can be verified that if we move probability $\varepsilon > 0$ from $q^{\bar{i}*}$ to q^{0*}, constraint (5.2) is relaxed while constraint (5.3) is unaffected. Thus the resulting buyer strategy is still implemented by π^*. The seller payoff achieved under the new buyer strategy is $W^* - \varepsilon\left(v^{\bar{i}} - c^{\bar{i}} - v^0\right)$. Thus, by varying ε, any payoff in the interval $\left[W^* - q^{\bar{i}}\left(v^{\bar{i}} - c^{\bar{i}} - v^0\right), W^*\right]$ is achievable. To achieve a payoff lower than $W^* - q^{\bar{i}}\left(v^{\bar{i}} - c^{\bar{i}} - v^0\right)$, we begin with the buyer strategy that has shifted the entire $q^{\bar{i}*}$ to q^{0*}, and induct the argument on the remaining highest investment action. At $q^0 = 1$, the payoff is v^0; thus any payoff in $[v^0, W^*]$ is achievable by π^*. $\square$

Proof of Proposition 5

Proof. Under the signal structure in (5.6):

$$\sum_{j\leq i} \pi\left(j|v^i\right)\left(v^i - v^j\right) - c^i = \pi\left(i-1|v^i\right)\left(v^i - v^{i-1}\right) - c^i = 0.$$

So (5.3) is always satisfied.[14] Next, a $\vec{q}$ and the signal structure in (5.6) jointly satisfy (5.2) if and only if:

$$\begin{aligned} & \left[q^i \pi\left(i|v^i\right) + q^{i+1}\pi\left(i|v^{i+1}\right)\right] v^i \geq q^{i+1}\pi\left(i|v^{i+1}\right) v^{i+1} && \forall i \leq m-1, \\ \iff \quad & q^i \geq \tfrac{\left[v^i - v^{i-1}\right]c^{i+1}}{\left[v^i - v^{i-1} - c^i\right]v^i} q^{i+1} && \forall i \leq m-1. \end{aligned}$$

Thus, (5.2) is satisfied under the signal structure in (5.6) if and only if $q \in \mathcal{Q}$.

Next, denote $\theta^i = \frac{\left(v^i - v^{i-1}\right)c^{i+1}}{\left(v^i - v^{i-1} - c^i\right)v^i}$ with the convention that $v^{-1} = 0$. So $\vec{q} \in \mathcal{Q}$ if and only if $q^i \geq \theta^i q^{i+1}$ $\forall i \leq m-1$. We claim that the highest seller payoff among all $\vec{q} \in Q$ is the $\vec{q}$ that satisfies $q^i = \theta^i q^{i+1}$ $\forall i \leq m-1$. To see this, suppose for a contradiction that $q^i > \theta^i q^{i+1}$. Denote $Q = q^i + q^{i+1}$, and denote $q^{i+1} = \lambda Q$ and $q^i = (1-\lambda) Q$. Thus $\frac{1-\lambda}{\lambda} > \theta^i$. Consider $\hat{\lambda} > \lambda$ such that $\frac{1-\hat{\lambda}}{\hat{\lambda}} = \theta^i$, and decrease q^i to $\hat{q}^i = \left(1-\hat{\lambda}\right) Q$ and increase q^{i+1} to $\hat{q}^{i+1} = \hat{\lambda} Q$. It is readily seen that condition (5.5) for all other $j \neq i$ will still be satisfied after this change, so this new buyer strategy is still in $\mathcal{Q}$. Since $v^{i+1} - v^i - c^{i+1} > 0$ $\forall i \leq m-1$, it implies that $v^{i+1} - c^{i+1} > v^i - c^i$, thus the change increases the payoff since the probability of taking the better investment action $i+1$ increases at the expense of the inferior action i. Therefore the highest seller payoff among all buyer strategy in $\mathcal{Q}$ is achieved under the strategy such that $q^i = \frac{1}{\theta^{i-1}} q^{i-1}$ $\forall i \geq 1$. We can write the probability of taking each investment action i as:

$$q^i = \frac{1}{\theta^{i-1}} q^{i-1} = \ldots = \left(\frac{1}{\theta^{i-1}} \times \frac{1}{\theta^{i-2}} \times \cdots \times \frac{1}{\theta^0}\right) q^0 = \underbrace{\prod_{j=0}^{i-1} \frac{\left(v^j - v^{j-1} - c^j\right) v^j}{\left(v^j - v^{j-1}\right) c^{j+1}}}_{\alpha^i} q^0$$

Note that $q^0 = 1 - \sum_{i=1}^m q^i = 1 - \sum_{i=1}^m \alpha^i q^0$; thus $q^0 = \frac{1}{1+\sum_{j=1}^m \alpha^j}$. Let $\alpha^0 = 1$ and we have $q^i = \frac{\alpha^i}{\sum_{j=0}^m \alpha^j}$ $\forall i$. The seller's payoff under this buyer strategy is thus $\sum_{i=0}^m q^i \left(v^i - c\right) = \sum_{i=0}^m \frac{\alpha^i}{\sum_{j=0}^m \alpha^j} \left(v^i - c^i\right)$. Following the argument behind Theorem 2.2, any seller payoff in the interval $\left[v^0, \sum_{i=0}^m \frac{\alpha^i}{\sum_{j=0}^m \alpha^j} \left(v^i - c^i\right)\right]$ is achievable by the signal structure in (5.6). $\square$

Proof of Proposition 6

Proof. As established in the main text, constraint (5.11) can be ignored. Since this is a linear program, the solution must lie on a vertex. It is readily verified that $q^2 = 0$ is never optimal under any γ. Moreover, by writing (5.10) as $q^1 \left(\frac{v^1 - v^0 - c^1}{v^1 - v^0}\right) \geq q^2 f(\gamma) \left(v^2 - v^1\right)$, if

[14] If $q^i = 0$, then $\pi\left(\cdot|v^i\right)$ does not affect the equilibrium conditions (5.2) and (5.3).

$q^1 = 0$, then $q^2 = 0$; thus $q^1 = 0$ also cannot be optimal. This implies that at least two of constraints (5.8) to (5.10) must bind at the solution.

Next, if $\left[q^1\left(\frac{c^1}{v^1-v^0}\right)+q^2\gamma\right]v^1 \geq q^2\gamma v^2 \iff \frac{q^1}{q^2} \geq \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$, then constraint (5.8) subsumes (5.9). On the other hand, constraint (5.10) requires that $\frac{q^1}{q^2} \geq \frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}}$. So if $\frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}} \geq \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}} \iff \gamma \leq \frac{c^1c^2}{(v^1-v^0)(v^2-v^1+c^1)}$, then constraint (5.9) is always subsumed by (5.8) and (5.10), so the solution is q^1 and q^2 such that (5.8) and (5.10) bind. Let:

$$\hat{\gamma} := \frac{c^1c^2}{\left(v^1-v^0\right)\left(v^2-v^1+c^1\right)}.$$

The following follows from the previous argument and some algebra:

Lemma 4. *When $\gamma \leq \hat{\gamma}$, the solution to program $\mathcal{P}(\gamma)$ is*

$$\begin{aligned} q^1 &= \left(\frac{v^0}{1+\left[c^1+v^0\right]h(\gamma)}\right)h(\gamma) \\ q^2 &= \left(\frac{v^0}{\left[1+\left[c^1+v^0\right]h(\gamma)\right]}\right)\frac{1}{\left[\gamma\left(v^1-v^0\right)+v^0\right]} \end{aligned}$$

where $h(\gamma) = \frac{\left[c^2-\gamma\left(v^2-v^0\right)\right]\left[v^1-v^0\right]}{\left[\gamma(v^1-v^0)+v^0\right]\left[v^1-v^0-c^1\right]v^1}$. The value is:

$$W(\gamma) = v^0 + \left(\frac{v^0}{\left[1+\left[c^1+v^0\right]h(\gamma)\right]}\right)\left[h(\gamma)\left(v^1-v^0-c^1\right)+\frac{v^2-v^0-c^2}{\gamma\left(v^1-v^0\right)+v^0}\right],$$

and

$$\frac{dW(\gamma)}{d\gamma} = \frac{v^0v^1\left(v^1-v^0\right)\left(v^1-v^0-c^1\right)\left[v^0v^1\left(v^1+c^2-c^1\right)-v^2\left(\left(v^0+c^1\right)\left(v^0+c^2\right)+v^1\left(v^1-c^1\right)\right)+\left(v^0+c^1\right)\left(v^2\right)^2\right]}{\left[c^1\left(v^0v^1+\gamma(v^1-v^0)(v^2+v^1-v^0)-c^2(v^1-v^0)\right)-(v^1-v^0)\left(v^0c^2+v^0v^1+\gamma\left((v^0)^2+(v^1)^2-v^0(v^1+v^2)\right)\right)\right]^2}.$$

Thus $\frac{dW(\gamma)}{d\gamma} \leq 0$ if and only if:

$$\begin{aligned} & v^0v^1\left(v^1+c^2-c^1\right)-v^2\left(\left(v^0+c^1\right)\left(v^0+c^2\right)+v^1\left(v^1-c^1\right)\right)+\left(v^0+c^1\right)\left(v^2\right)^2 \leq 0 \\ \iff \quad & \frac{v^2-v^0-c^2}{v^1-v^0-c^1} \leq \frac{c^2\left(v^1-v^0\right)+v^0\left(v^2-v^0\right)}{(c^1+v^0)(v^2-v^0)-(v^1-v^0-c^1)v^1} = \xi. \end{aligned}$$

Next, consider $\gamma > \hat{\gamma}$. Suppose for a contradiction that constraint (5.9) does not bind; then this implies that constraints (5.8) and (5.10) must bind. The binding (5.10) implies that $\frac{q^1}{q^2} = \frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}} < \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$ where the inequality follows from $\gamma > \hat{\gamma}$. This implies that $\left[q^1\left(\frac{c^1}{v^1-v^0}\right)+q^2\gamma\right]v^1 < q^2\gamma v^2$, so if constraint (5.8) binds, then constraint (5.9) must be violated. Contradiction.

Therefore, if $\gamma > \hat{\gamma}$, then constraint (5.9) is binding under the solution of program $\mathcal{P}(\gamma)$. There are then two cases to consider:

Case A: Constraints (5.8) and (5.9) bind. Let $\left(q_A^1(\gamma), q_A^2(\gamma)\right)$ be the solution to the system of equations whereby (5.8) and (5.9) hold with equality. Solving it, we have:

$$\begin{aligned} q_A^1(\gamma) =& \frac{\gamma v^0\left(v^1-v^0\right)\left(v^2-v^1\right)}{\gamma v^2\left(c^1+v^0\right)\left(v^1-v^0\right)+v^0v^1\left(c^1-\gamma\left(v^1-v^0\right)\right)} \\ q_A^2(\gamma) =& \frac{v^0v^1c^1}{\gamma v^2\left(c^1+v^0\right)\left(v^1-v^0\right)+v^0v^1\left(c^1-\gamma\left(v^1-v^0\right)\right)} \end{aligned}$$

Under $\left(q_A^1(\gamma), q_A^2(\gamma)\right)$, the value is:

$$\begin{aligned} W_A(\gamma) :=& v^0+q_A^1(\gamma)\left(v^1-v^0-c^1\right)+q_A^2(\gamma)\left(v^2-v^0-c^2\right) \\ =& \frac{v^0v^1\left[\gamma\left(v^1-v^0\right)\left(v^2-v^1\right)+c^1\left(v^2-c^2+\gamma\left(v^1-v^0\right)\right)\right]}{\gamma v^2\left(v^0+c^1\right)\left(v^1-v^0\right)-v^0v^1\left(\gamma\left(v^1-v^0\right)-c^1\right)}, \end{aligned}$$

and it can be verified that $W_A(\gamma)$ is strictly decreasing:

$$\begin{aligned} \frac{dW_A(\gamma)}{d\gamma} =& -\frac{c^1v^0v^1\left(v^1-v^0\right)\left[v^0\left(v^2-v^1\right)\left(v^2-v^1-c^2\right)+c^1\left(v^2\left(v^2-c^2\right)-v^0v^1\right)\right]}{\left[v^0\left(c^1+\gamma\left(v^0-v^1\right)\right)v^1-\gamma\left(c^1+v^0\right)\left(v^0-v^1\right)v^2\right]^2} \\ <& -\frac{c^1v^0v^1\left(v^1-v^0\right)\left[v^0\left(v^2-v^1\right)\left(v^2-v^1-c^2\right)+c^1v^2\left(v^2-v^0-c^2\right)\right]}{\left[v^0\left(c^1+\gamma\left(v^0-v^1\right)\right)v^1-\gamma\left(c^1+v^0\right)\left(v^0-v^1\right)v^2\right]^2} < 0 \end{aligned}$$

Case B: (5.9) and (5.10) bind. Let $\left(q_B^1(\gamma), q_B^2(\gamma)\right)$ be the solution to the system of equations whereby (5.9) and (5.10) hold with equality. Solving it, we have:

$$\begin{aligned} q_B^1(\gamma) =& \frac{v^0\left(v^1-v^0\right)\left(c^2+\gamma\left(v^2-v^0\right)\right)}{\left(v^1-v^0-c^1\right)\left[v^0c^2+v^0v^1+\gamma\left(v^1-v^0\right)\left(v^2-v^0\right)\right]} \\ q_B^2(\gamma) =& \frac{v^0v^1}{v^0c^2+v^0v^1+\gamma\left(v^1-v^0\right)\left(v^2-v^0\right)} \end{aligned}$$

Under $\left(q_B^1(\gamma), q_B^2(\gamma)\right)$, the value is:

$$\begin{aligned} W_B(\gamma) :=& v^0 + q_B^1(\gamma)\left(v^1 - v^0 - c^1\right) + q_B^2(\gamma)\left(v^2 - v^0 - c^2\right) \\ =& \frac{v^0 v^1 v^2}{c^2 v^0 + v^0 v^1 + \gamma\left(v^1 - v^0\right)\left(v^2 - v^0\right)}; \end{aligned}$$

It is immediate that $W_B(\gamma)$ is strictly decreasing.

If constraints (5.8) and (5.9) bind, then $\left[q^1\left(\frac{c^1}{v^1-v^0}\right) + q^2\gamma\right] v^1 = q^2\gamma v^2 \iff \frac{q^1}{q^2} = \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}} > \frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}}$, where the last inequality follows from $\gamma > \hat{\gamma}$. The last inequality would imply constraint (5.10).

If, instead, constraints (5.9) and (5.10) bind, then the binding (5.10) implies that $\frac{q^1}{q^2} = \frac{f(\gamma)\left(v^2-v^1\right)}{\frac{v^1-v^0-c^1}{v^1-v^0}} < \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$. The last inequality $\frac{q^1}{q^2} < \frac{\gamma\left(v^2-v^1\right)}{\frac{c^1}{v^1-v^0}}$ implies that $\left[q^1\left(\frac{c^1}{v^1-v^0}\right) + q^2\gamma\right] v^1 < q^2\gamma v^2$ which means that constraint (5.8) is satisfied.

Therefore, for any $\gamma > \hat{\gamma}$, both $W_A(\gamma)$ and $W_B(\gamma)$ are attainable. The following result then follows:

Lemma 5. *When $\gamma > \hat{\gamma}$, the value function of program $\mathcal{P}(\gamma)$ is $W(\gamma) = \max\left\{W_A(\gamma), W_B(\gamma)\right\}$. Since both W_A and W_B are strictly decreasing, W is also strictly decreasing.*

The proposition then follows from Lemma 4 and 5. □

References

Aumann, Robert J and Michael Maschler. 1995. *Repeated Games with Incomplete Information.* MIT press.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The Limits of Price Discrimination." *American Economic Review* 105 (3):921–57.

Condorelli, Daniele and Balazs Szentes. 2017. "Information Design in the Hold-up Problem." .

Gibbons, Robert. 1992. *Game Theory for Applied Economists.* Princeton University Press.

González, Patrick. 2004. "Investment and Screening under Asymmetric Endogenous Information." *The RAND Journal of Economics* :502–519.

Gul, Faruk. 2001. "Unobservable Investment and the Hold-Up Problem." *Econometrica* 69 (2):343–376.

Halac, Marina. 2015. "Investing in a relationship." *The RAND Journal of Economics* 46 (1):165–185.

Hermalin, Benjamin E. 2013. "Unobserved Investment, Endogenous Quality, and Trade." *The RAND Journal of Economics* 44 (1):33–55.

Hermalin, Benjamin E. and Michael L. Katz. 2009. "Information and the Hold-up Problem." *The RAND Journal of Economics* 40 (3):405–423.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101 (6):2590–2615.

Lau, Stephanie. 2008. "Information and Bargaining in the Hold-up Problem." *The RAND Journal of Economics* 39 (1):266–282.

Ostrovsky, Michael and Michael Schwarz. 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics* 2 (2):34–63.

Rayo, Luis and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of Political Economy* 118 (5):949–987.

Roesler, Anne-Katrin and Balázs Szentes. 2017. "Buyer-Optimal Learning and Monopoly Pricing." *American Economic Review* 107 (7):2072–80.

Tan, Teck Yong. 2017. "The Extrinsic Motivation of Freedom at Work." .